PREDICTIVE MAINTENANCE OF ESSENTIAL MACHINERY: A DATA-DRIVEN APPROACH FOR OIL AND GAS INDUSTRY RESILIENCE

Team 10

Chia Wei Kit, Samuel

Dhruval Kenal Kothari

Lim En Ning

Riley Ang Xile

Tan Zheng Da Jared

Jared

Opportunity Statement

A key business problem for Aramco is to optimize the maintenance and performance of its critical assets, such as refinery equipment and machinery. The opportunity lies in implementing predictive maintenance using advanced analytics and machine learning techniques.

Rising Cost of Unplanned Downtime

According to Siemens' report in 2023,

1. The cost of one hour of downtime has more than doubled from 2021 to 2022 to nearly $500,000
2. Loss of $1.5 Trillion annually for Fortune Global 500 companies

Jared

Saudi Aramco

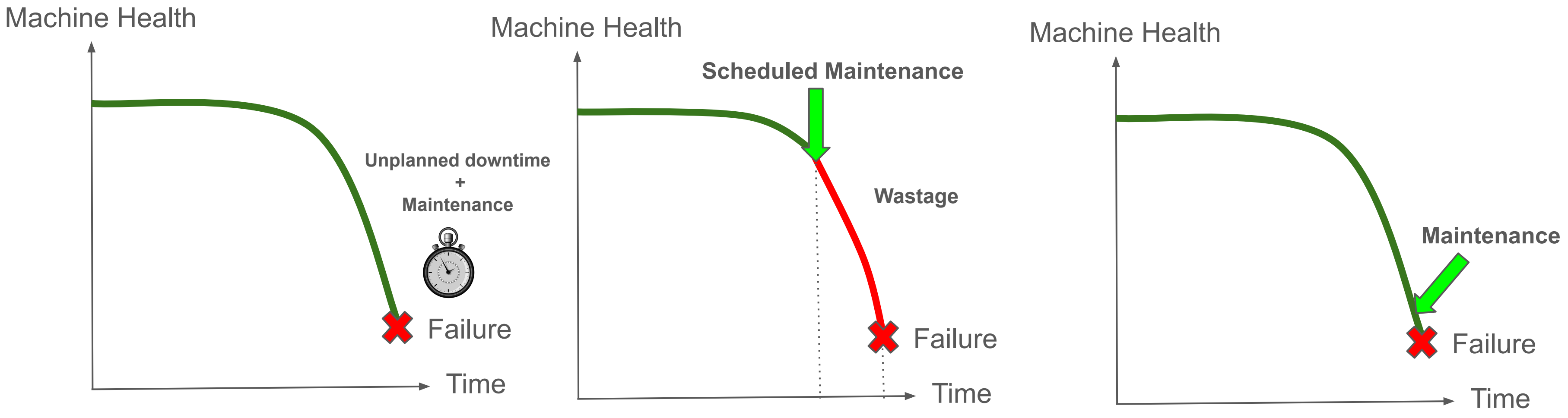
Handles 12.8 million barrels of oil equivalent per day,

Unplanned downtime would cause:

1. **Significant Lost in Profits**
2. **Potential Safety Incidents**

Jared

3 Solutions to Deal with Machine Failure



Reactive

Preventive

Predictive

Jared

NASA Turbofan Engine Degradation Simulation Data Set

Allows for exploration and modelling of engine degradation patterns

Predict Remaining Useful Life of critical and non-critical machinery

Improve timing of maintenance, avoid unplanned downtime & production disruptions

Increase profitability, reduce safety incidents, and improve overall competitiveness in the industry

What is currently being done by others and how can Aramco learn

Corrective Maintenance | “Run to failure”

- High costs and longer turnaround time

Anomaly Detection

- Analyse data points
- Find outliers or anomalies
- Detect them as needs for maintenance

Prognostics and Health Management

- Is a predictive maintenance strategy
- Forecasting equipment failures by predicting downtimes
- Based on data collected and historical trends

PEPSICO

Downtime of just 2.88%

noranda alumina

60% decline in bearing changes at the plant

SMRT

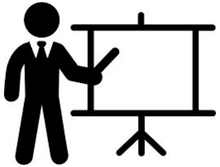
Record 1 million kilometers between failure in Singapore trains

Dhruval

What are some gaps in the current predictive maintenance solutions

Too much data available which might not be compatible

Financial benefits vs costs of implementing sensors



Workforce training

Poor quality of data and need to properly handle missing data

En Ning

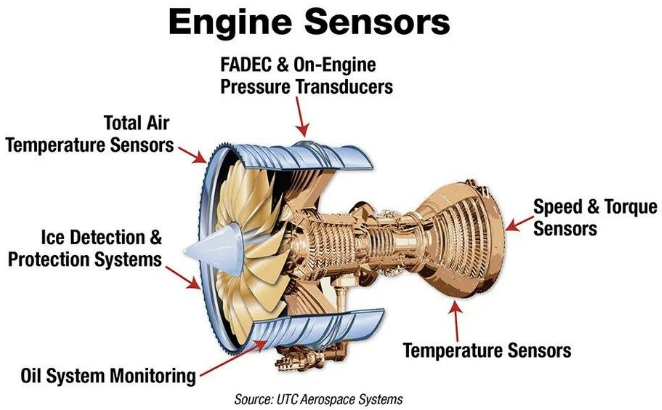
Modelling of Run-to-Failure Dataset using C-MAPSS

En Ning

Modelling of Run-to-Failure Dataset

Damage modelling reflects real-world operationality of aircraft engines to estimate the degree and rate of degradation after each flight.

Same engine used under various flight and operational conditions



Sensors are used to collect data on each flight to reflect engine degradation

Modelling of Run-to-Failure Dataset

Damage modelling reflects real-world operationality of aircraft engines to estimate the degree and rate of degradation after each flight.

Initial wear & tear from manufacturing and transportation of engines

Impact of maintenance work done on engines between flights

En Ning

Remaining Useful Life (RUL)

RUL is used to predict the number of operational cycles an engine can run before failure occurs.

Discern the best time to **pause production and conduct maintenance**.

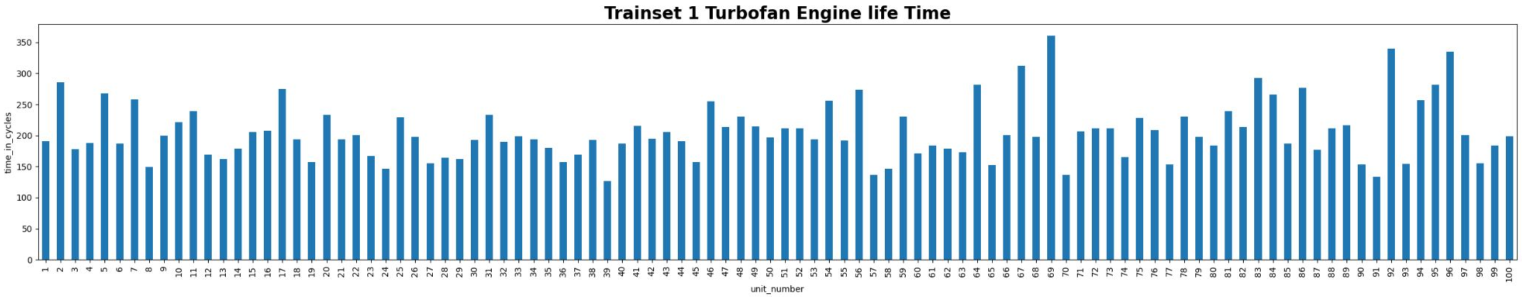
Prevents **catastrophic failure** if maintenance is not done on time.

Prevents **downtime** of aircraft if maintenance is not done on time.

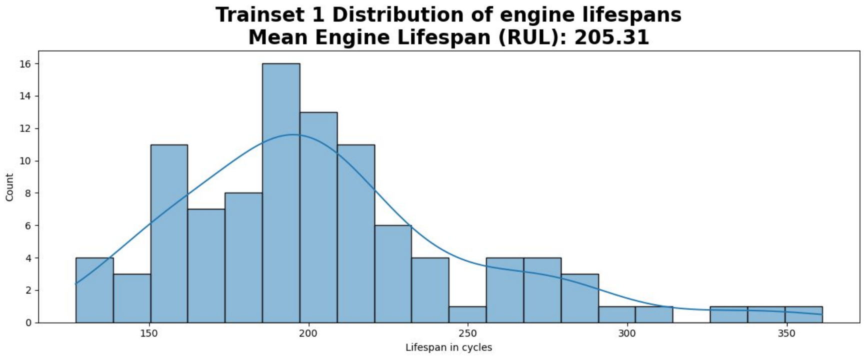
En Ning

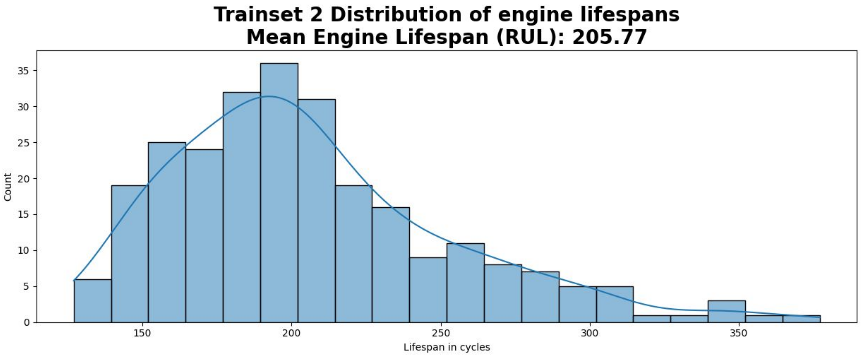
Remaining Useful Life (RUL)

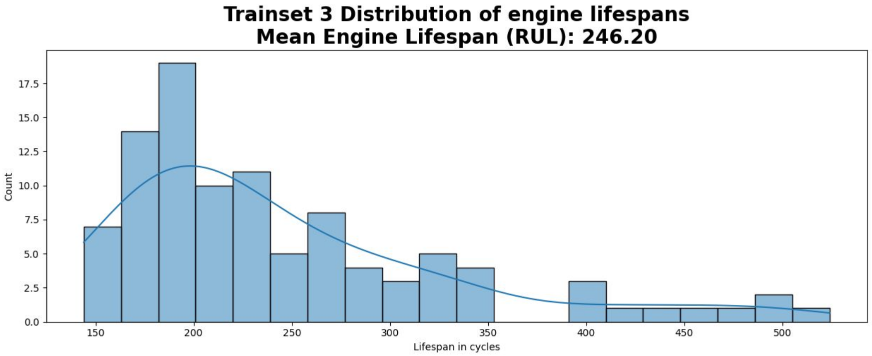
Though each fleet of engines **operates under the same conditions** and are of the **same type**, each engine starts with <u>different degrees of wear and tear, and natural manufacturing variation</u>.

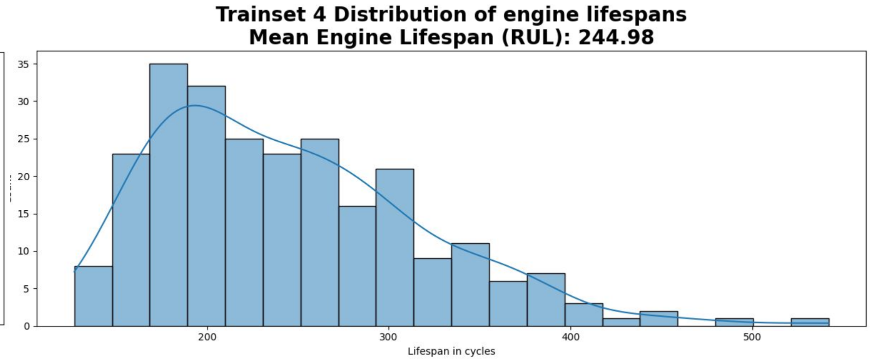


RUL: Distribution of Engine Lifespans (Train Set)





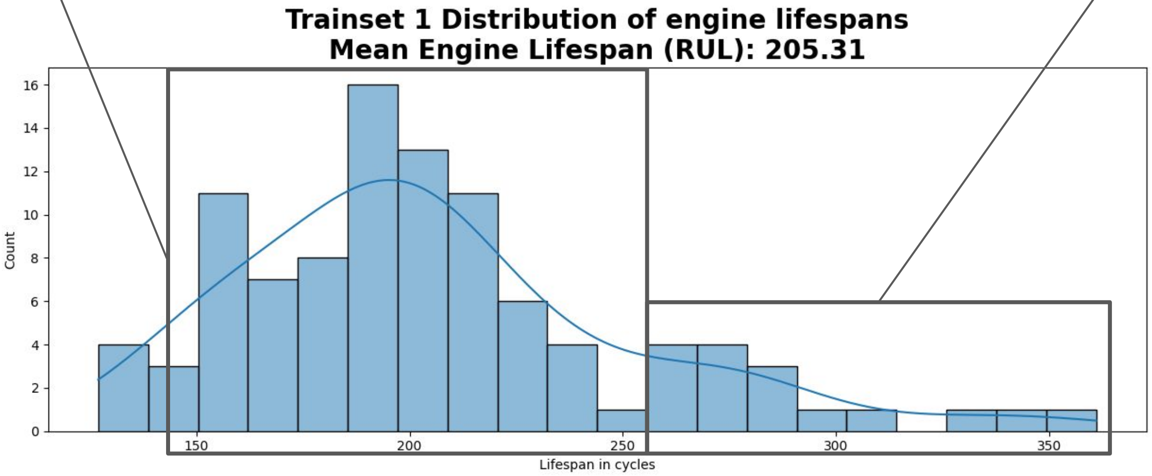




RUL: Distribution of Engine Lifespans (Train Set)

Most engines in the fleet have lifespans of approximately 150-250 cycles

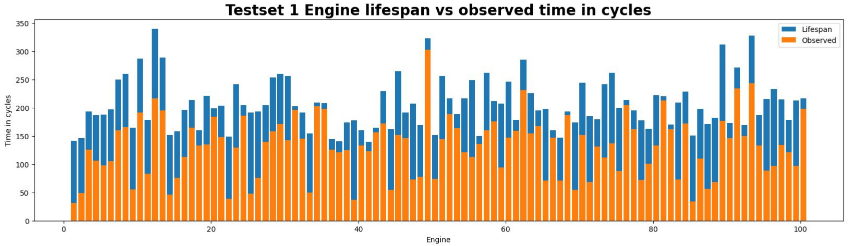
A few outliers significantly exceeds the average lifespan.

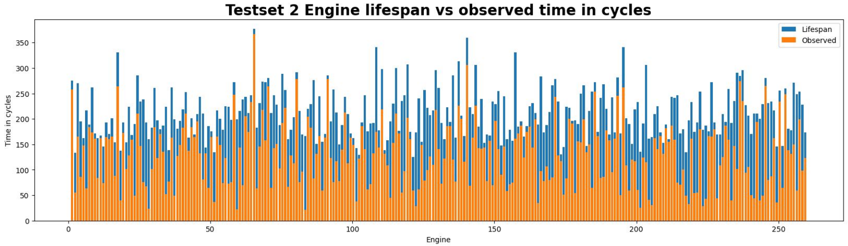


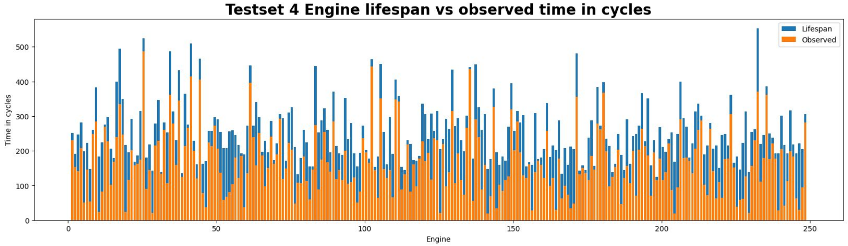
This is a typical distribution for reliability and life data analysis where the engine is subjected to various failure mechanisms, resulting in wide range of life spans.

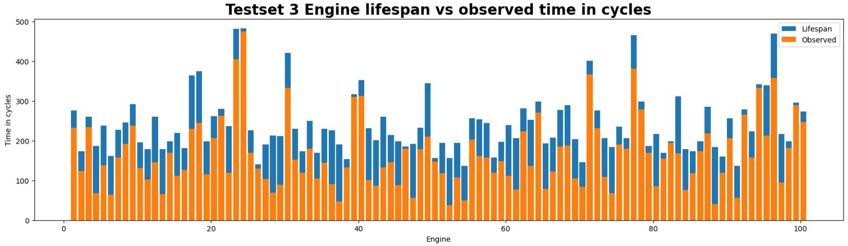
En Ning

RUL: Engine Lifespan vs Observed Time (Test Set)







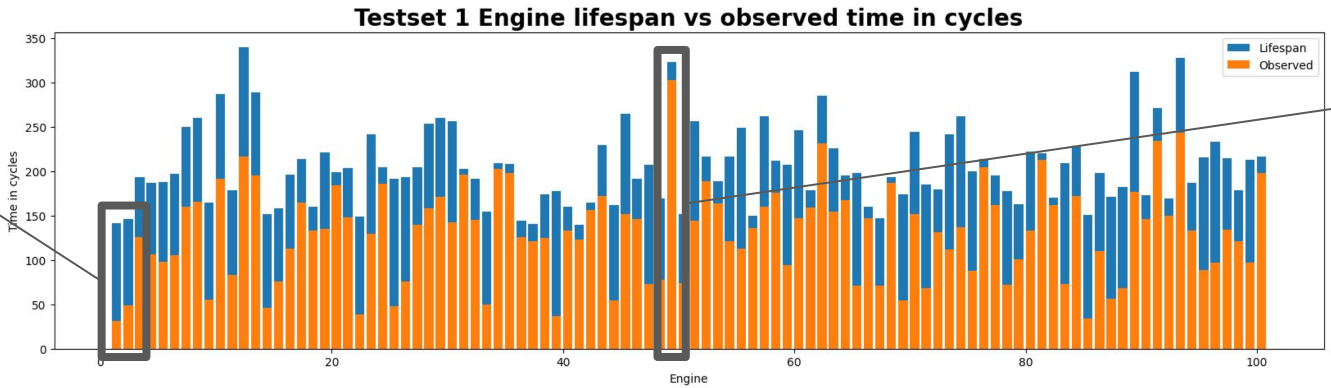


En Ning

RUL: Engine Lifespan vs Observed Time (Test Set)

Observed operational life is a proportion of the actual full lifespans of engines.

Fails **long after** observed operational life.



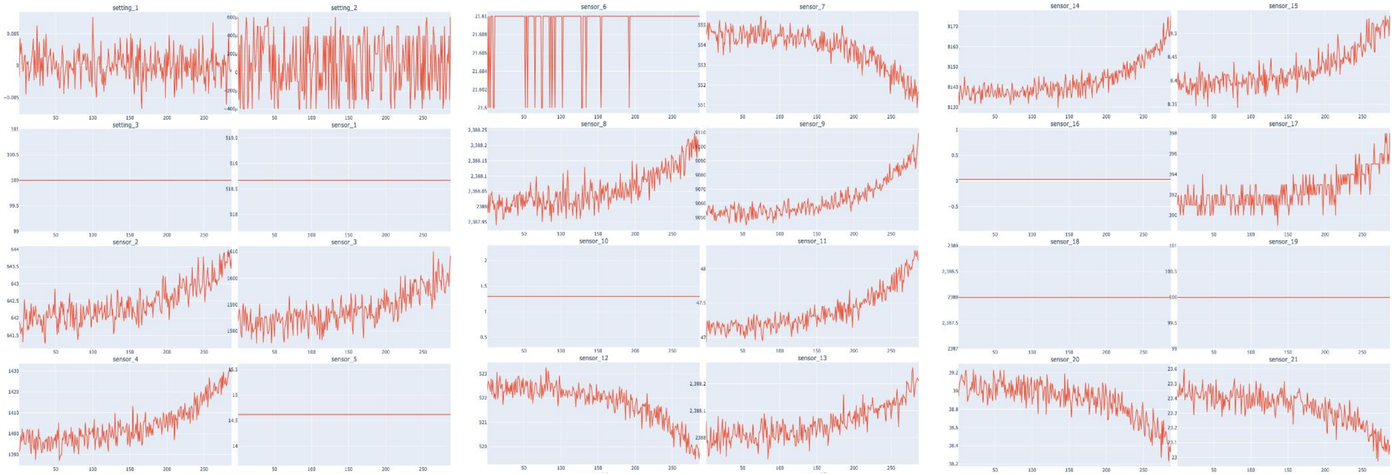
Fails **soon after** observed operational life.

The test set gives a challenging variety of scenarios where some engines fail soon after we observe them while others fail long after we observe them.

En Ning

Sensor and Setting Data

FD001 Training Engine #1



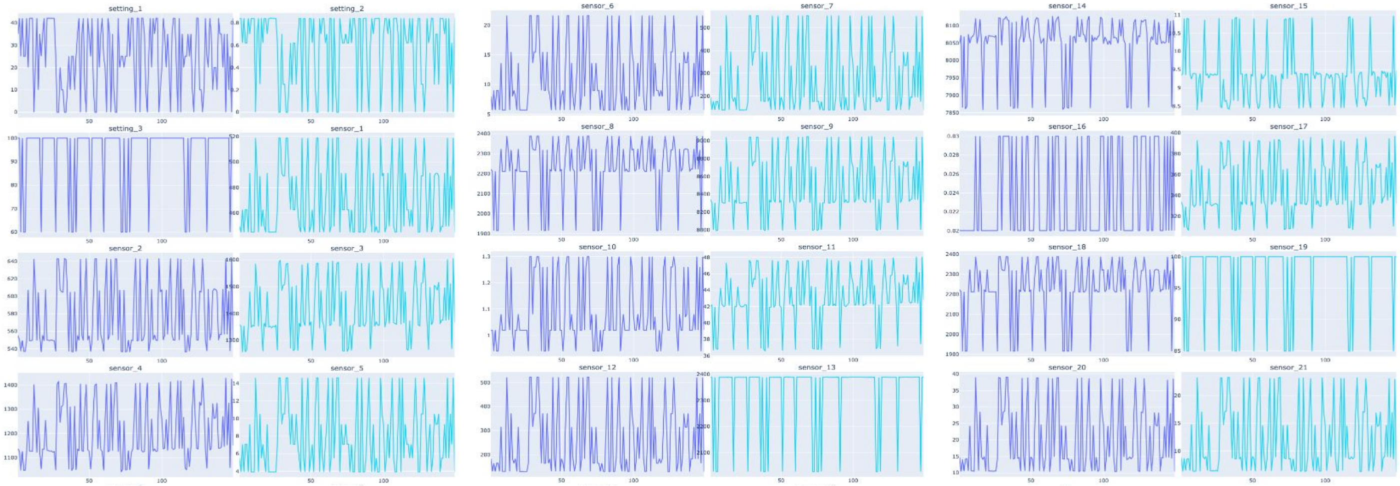
Each engine has 3 operational settings and readings from 21 different sensors. Each data point captures a snapshot of the **engine's performance during a single operational cycle**.

En Ning

Sensor and Setting Data

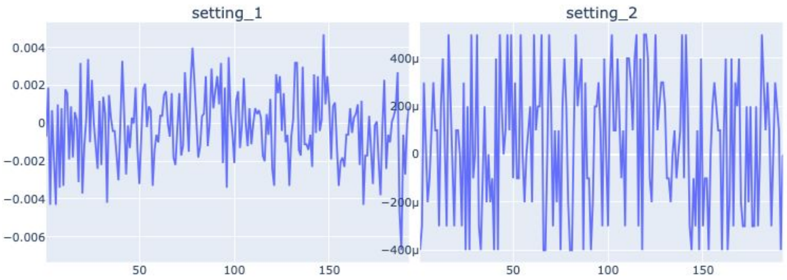
Engine degradation is simulated through **changes in operational settings and sensor readings**, thus each fleet has significantly different data.

FD002 Training Engine #1

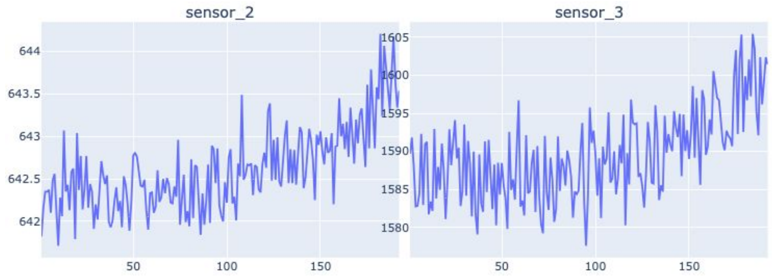


This temporal information constitutes the **independent variables** in our analysis for the **prediction of RUL** of each engine.

Settings and Sensor Data



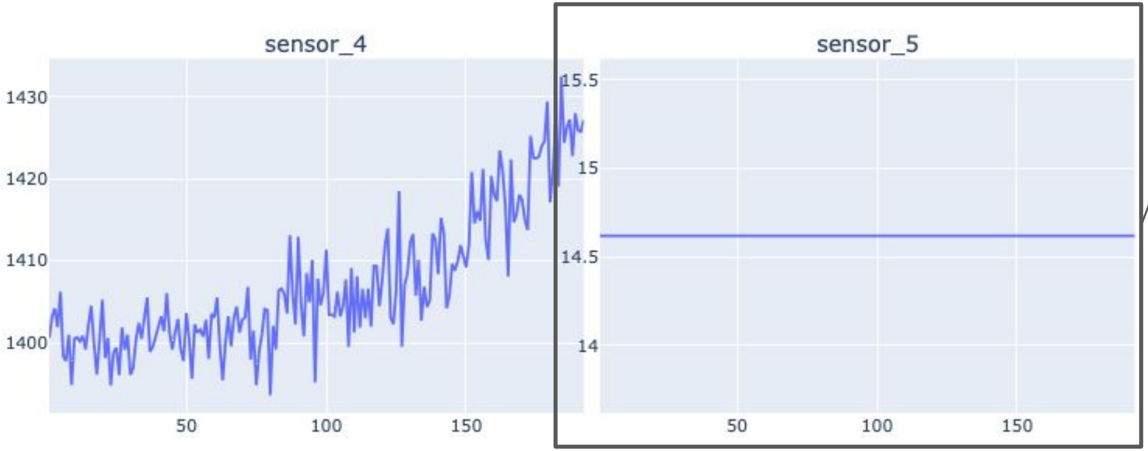
Operational settings are **changed between a fixed range of values**, as expected of setting values.



Sensor values vary around a mean and eventually **trend exponentially higher** as the **engine reaches system failure**.

Noisy Data

Reading of sensor 5 remains constant throughout the whole operational life.

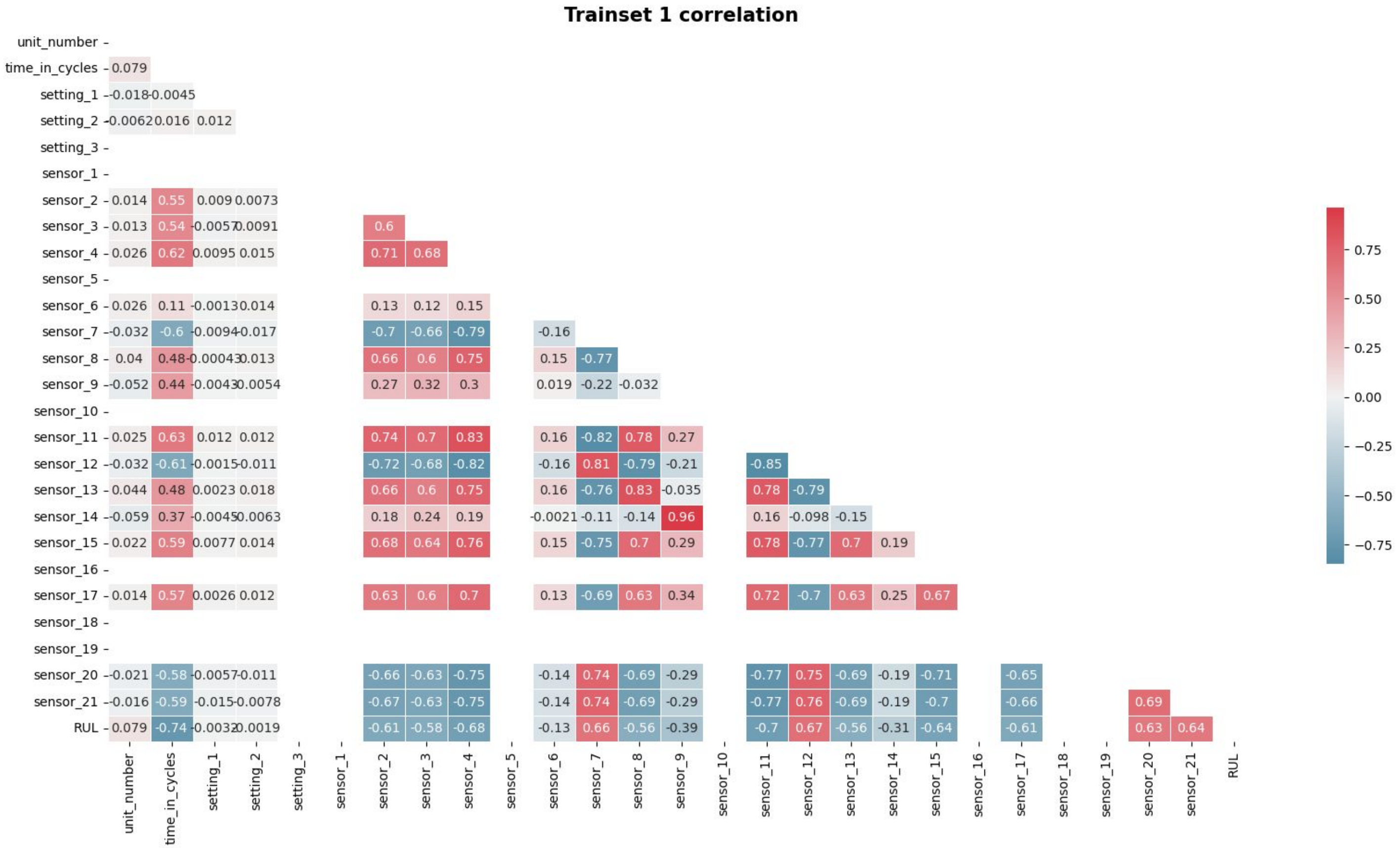


Sensor 5's readings are irrelevant in predicting RUL and can be removed.
Other sensor measurements that contain noisy data will be removed too.

Feature Engineering & Feature Selection

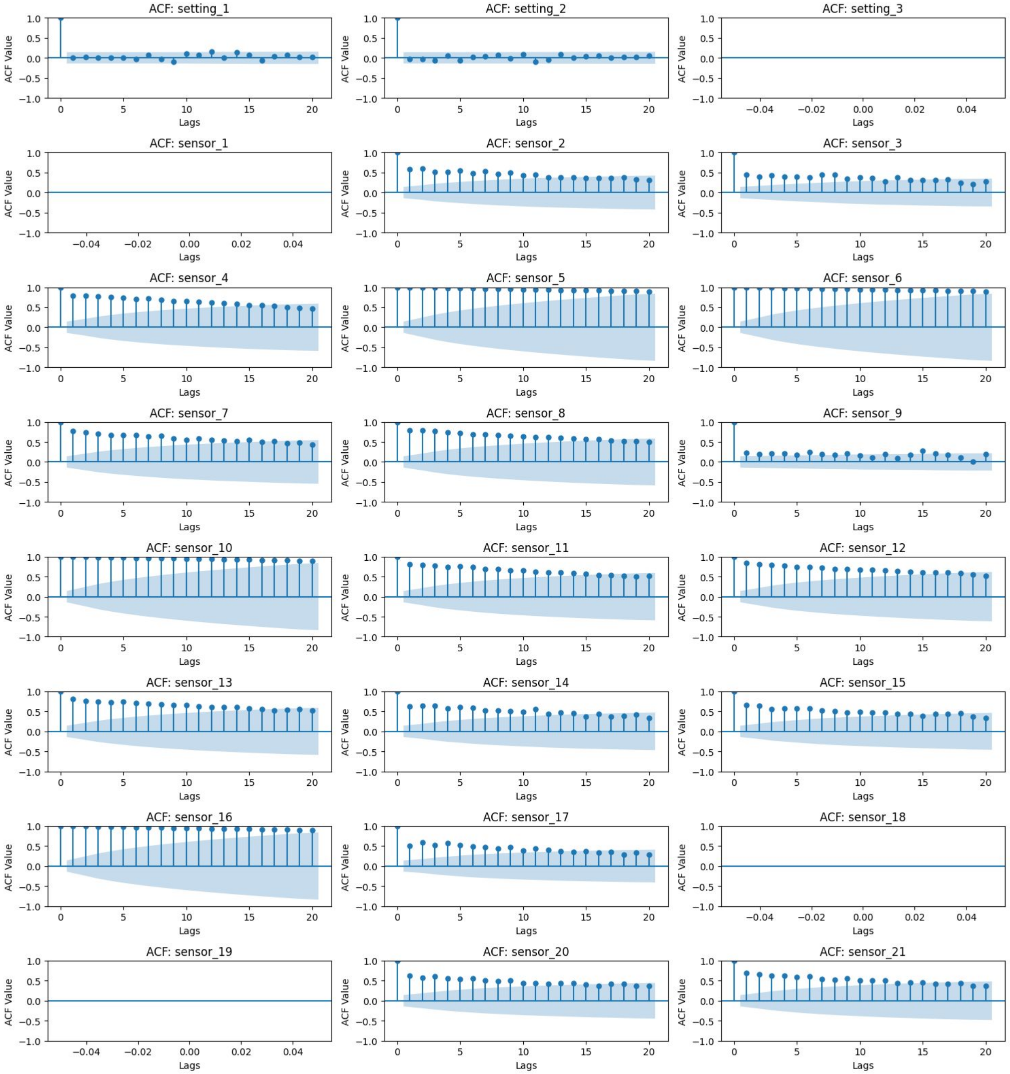
Samuel

Correlation



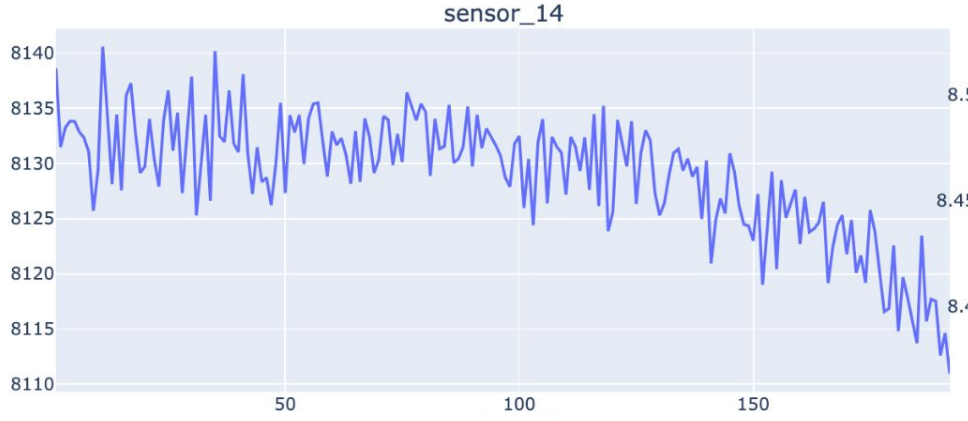
Samuel

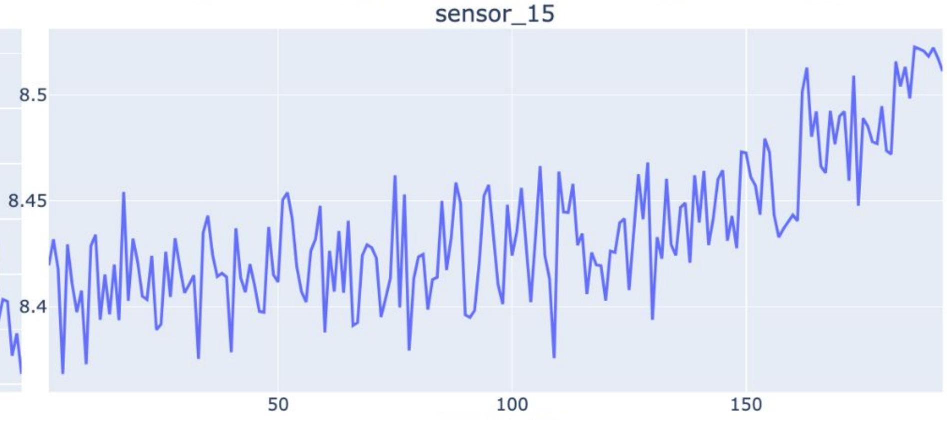
Autocorrelation

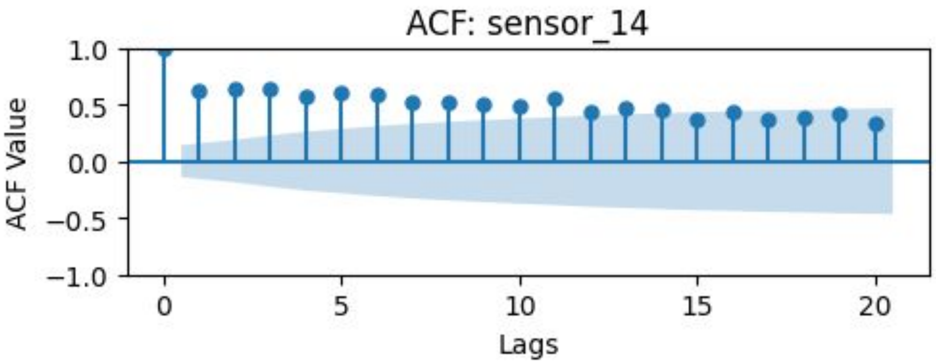


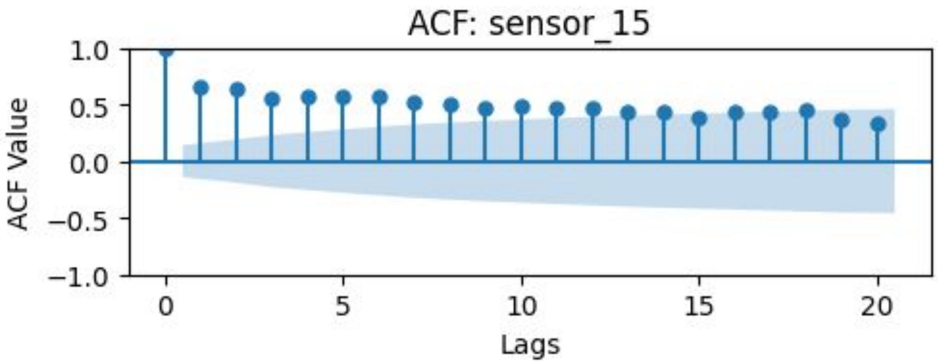
Samuel

Autocorrelation





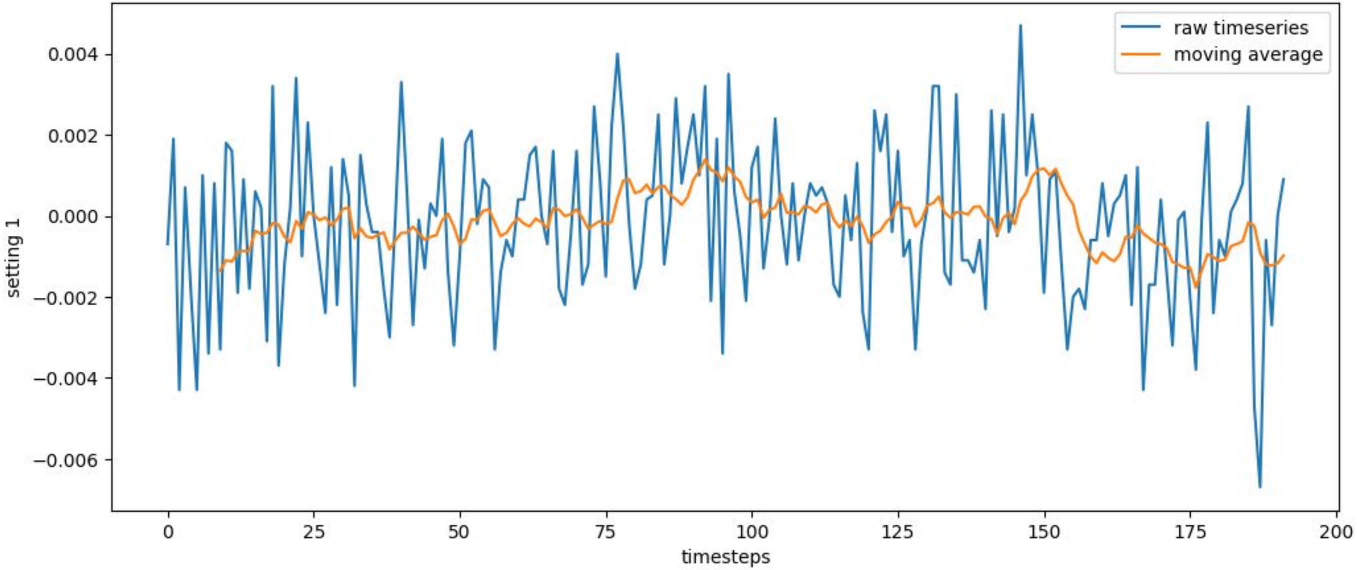




Samuel

Feature Engineering – Temporal Feature Creation

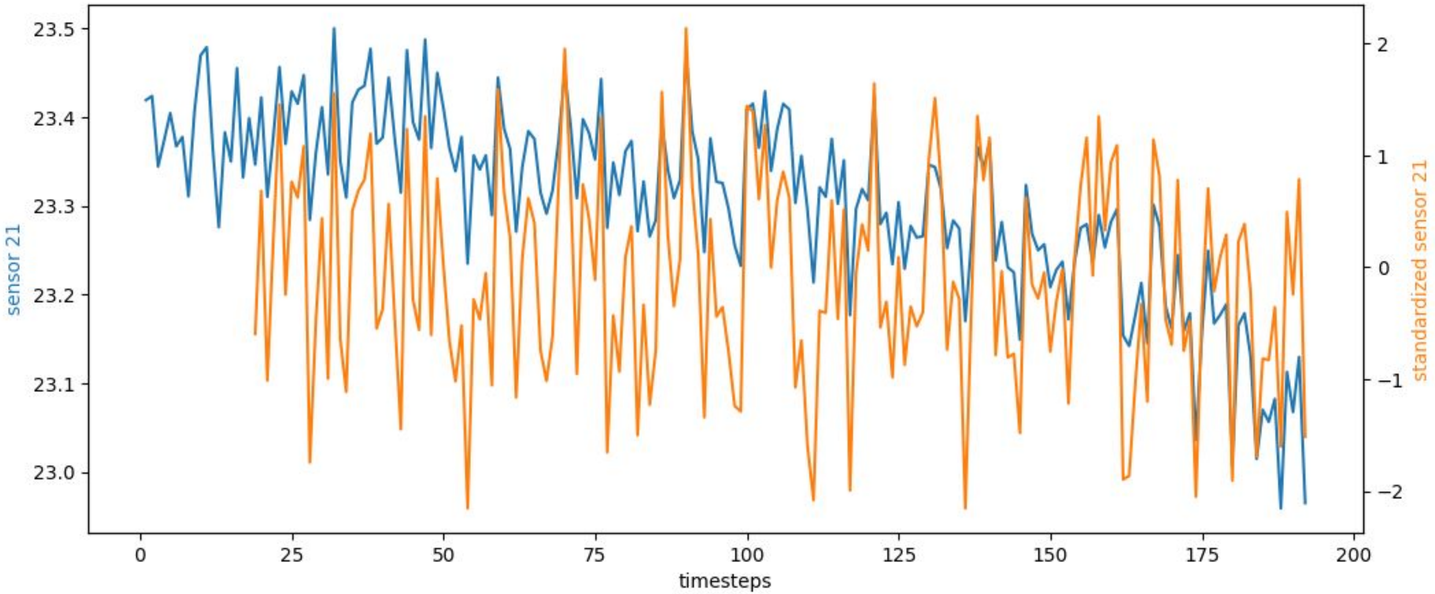
- ACF and PACF plots show that temporal effects are present
- Moving averages can help capture the trend and smooth out noisy time series



Samuel

Feature Engineering – Standardization

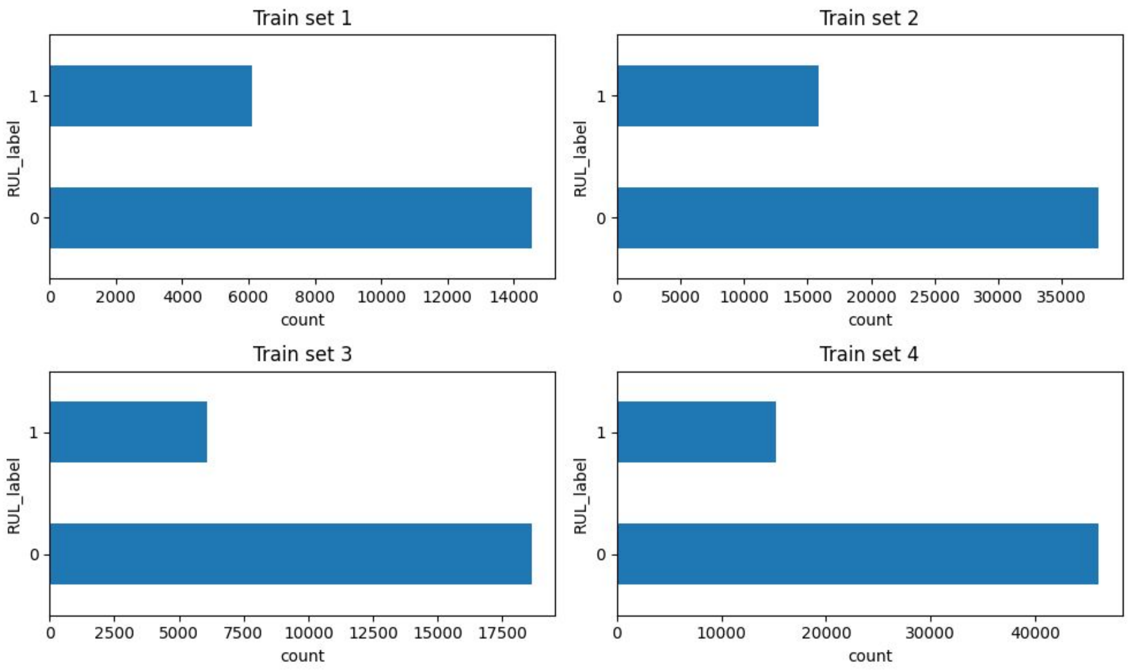
- Prevent models from having a bias towards features with higher magnitude
- Avoid look-ahead bias with rolling z-score



Samuel

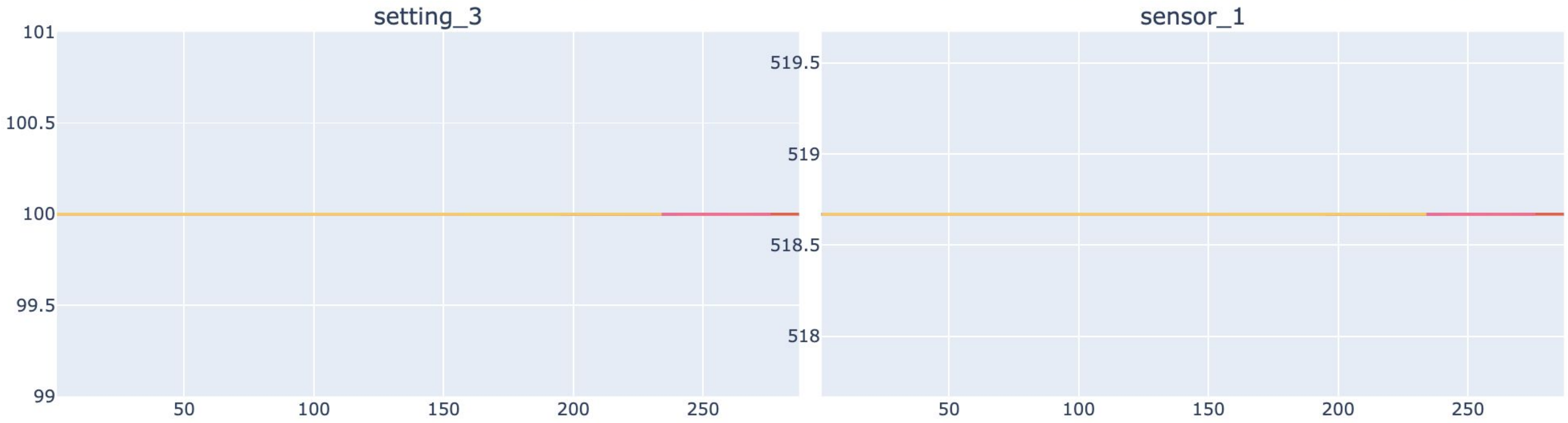
Feature Engineering – Classification Labels

- Adding labels for target variables
- 0 for ‘no risk’, RUL > 60
- 1 for ‘at risk’, RUL <= 60
- Unbalanced datasets

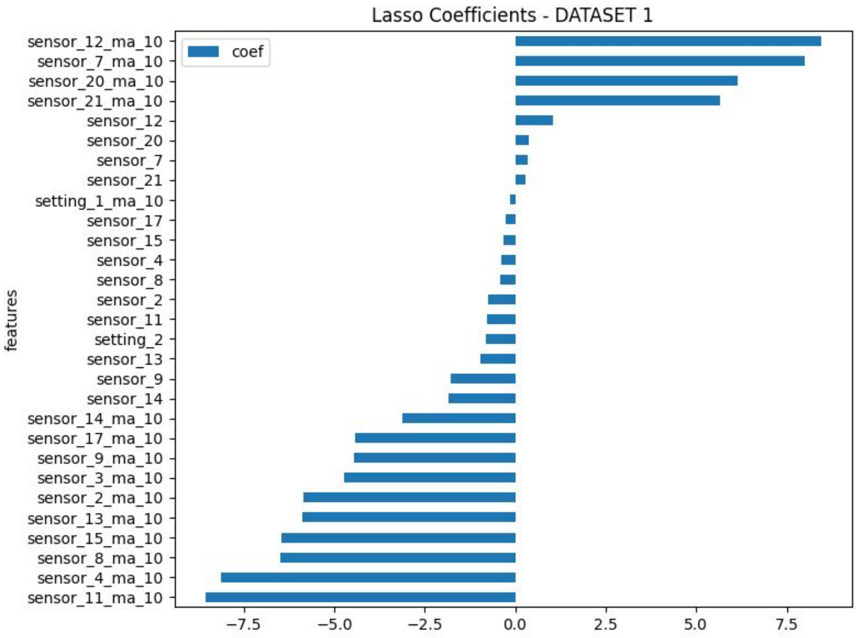


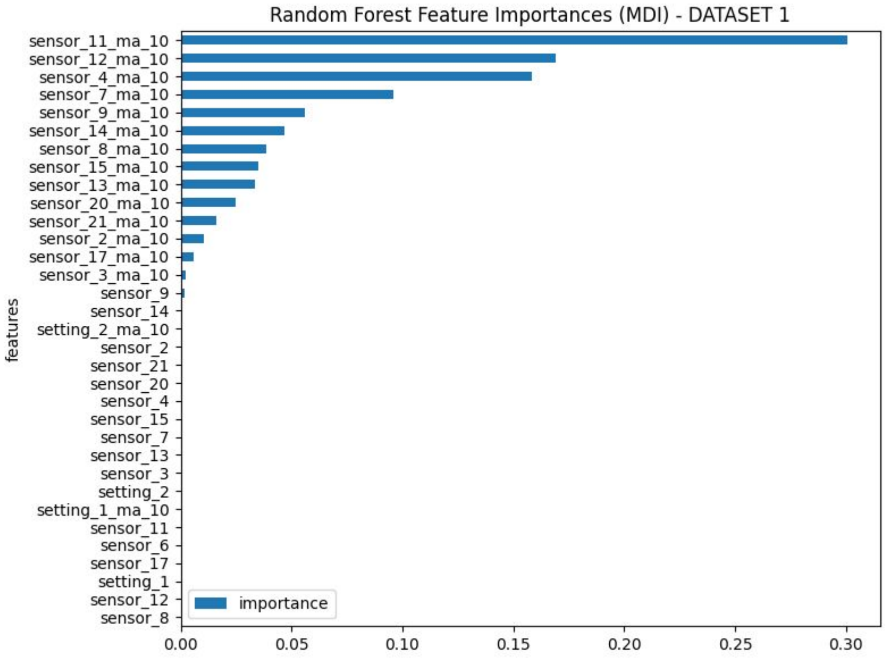
Feature Selection – Irrelevant Features Elimination

- Removed any features with constant values



Feature Selection





Regression & Classification Models

Modelling

Regression Models

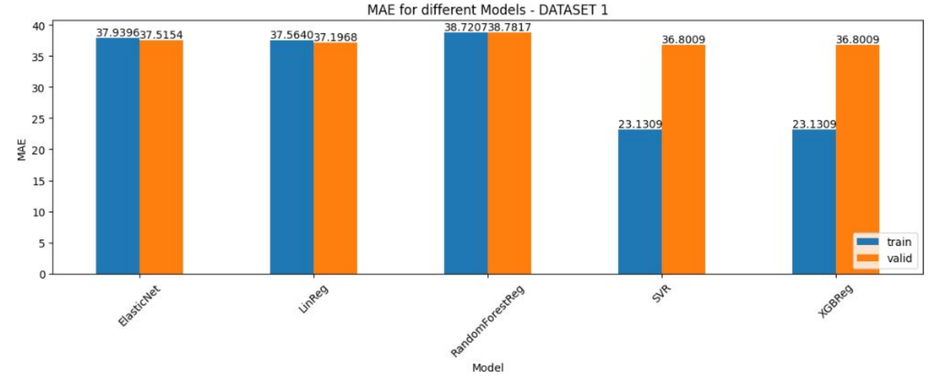
1. Linear Regression
2. Elastic Net
3. Random Forest Regressor
4. XGBoost Regressor
5. Support Vector Regression

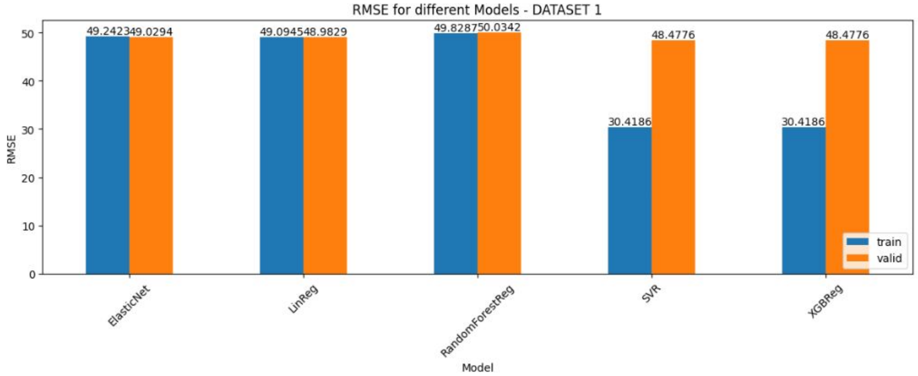
Classification Models

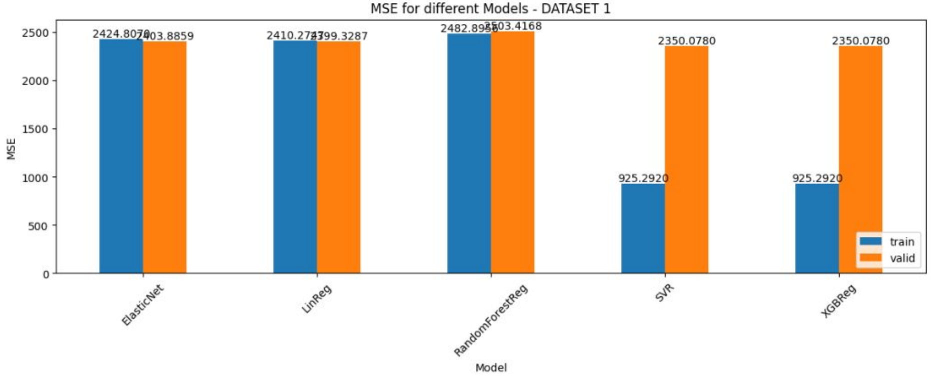
1. Gaussian Naive Bayes
2. Decision Trees
3. Random Forest Classifier
4. XGBoost Classifier
5. K-Nearest Neighbors

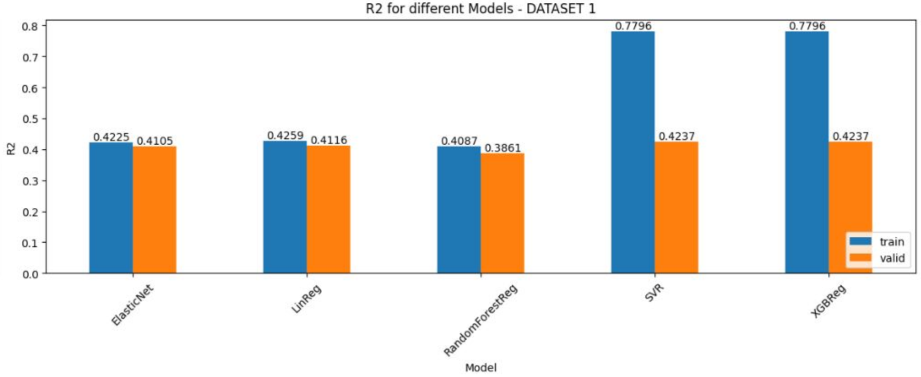
Riley

Regression Models

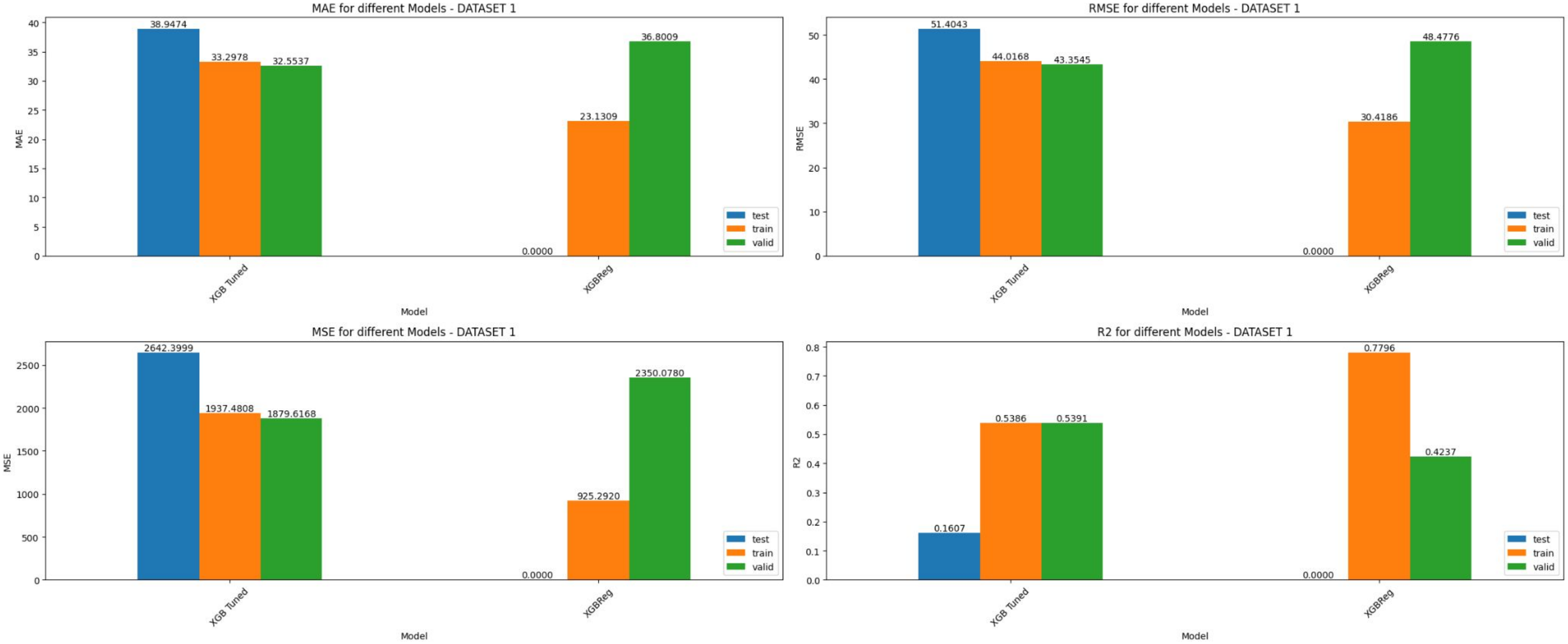




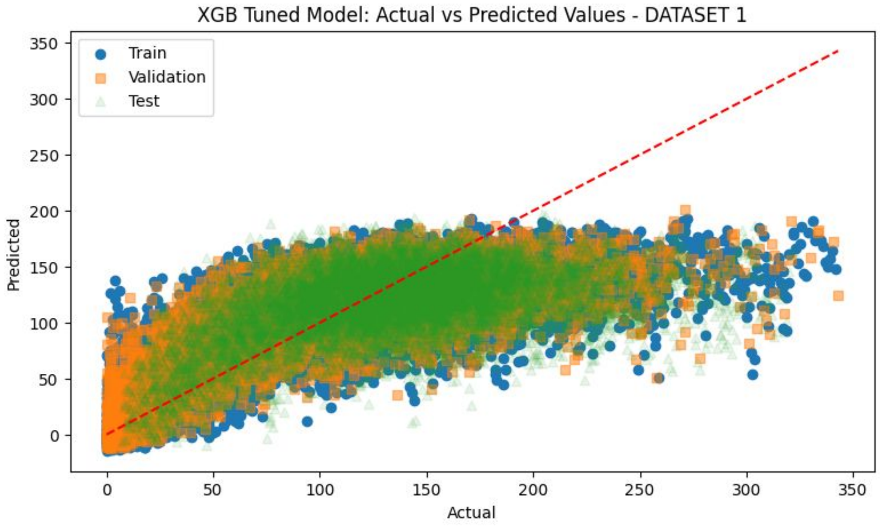


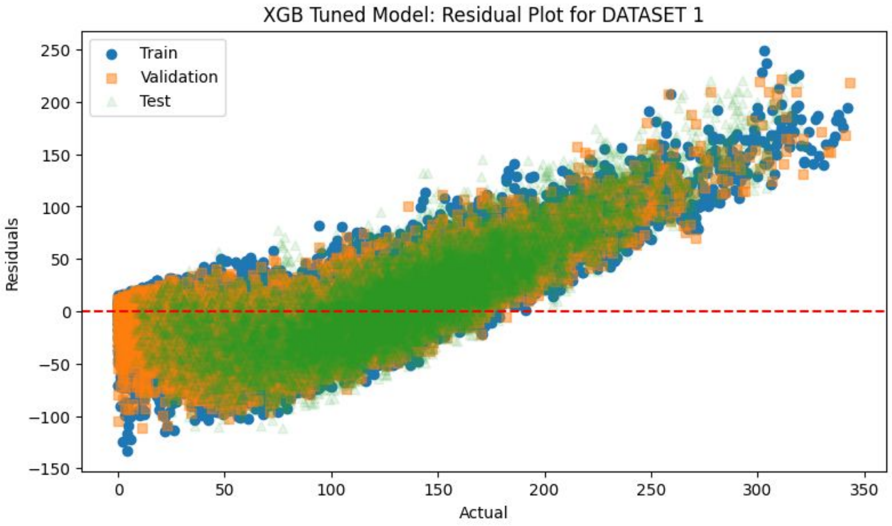


Regression Models – Hyperparameter Tuning

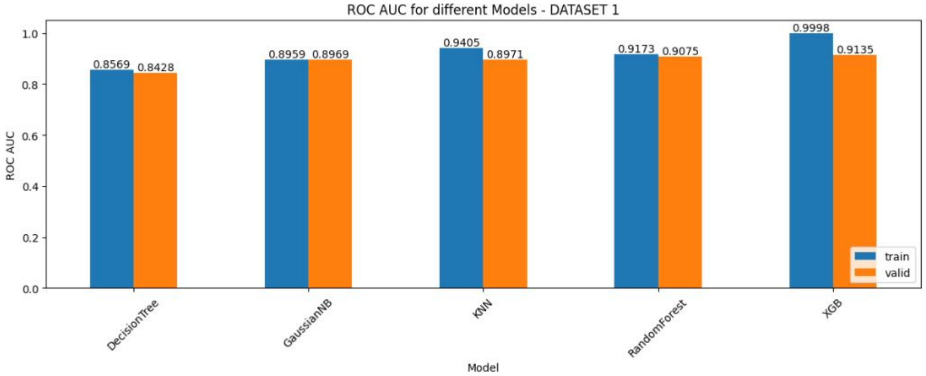


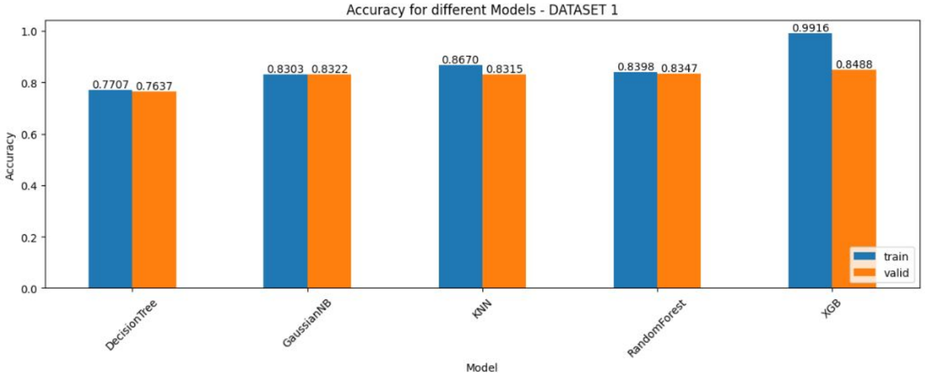
Regression Models – Hyperparameter Tuning

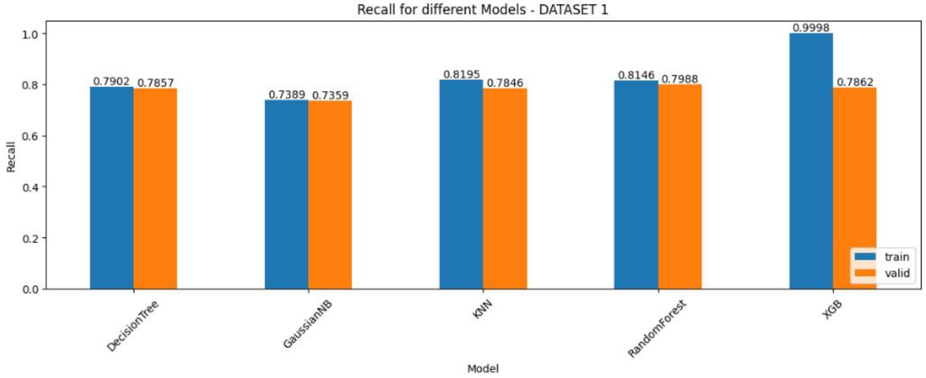


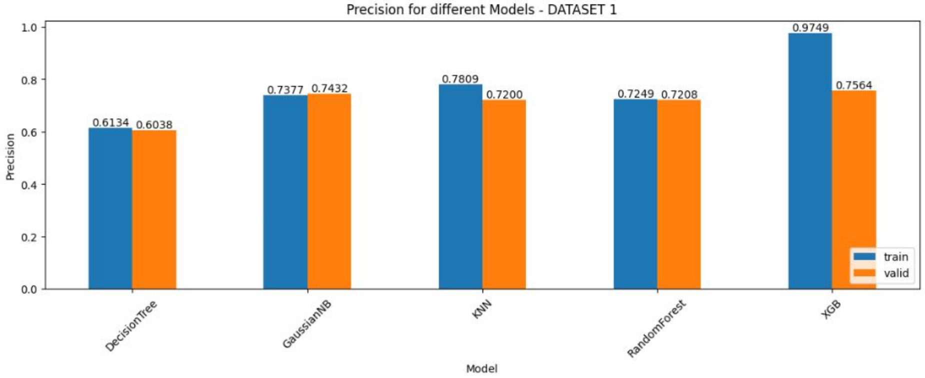


Classification Models



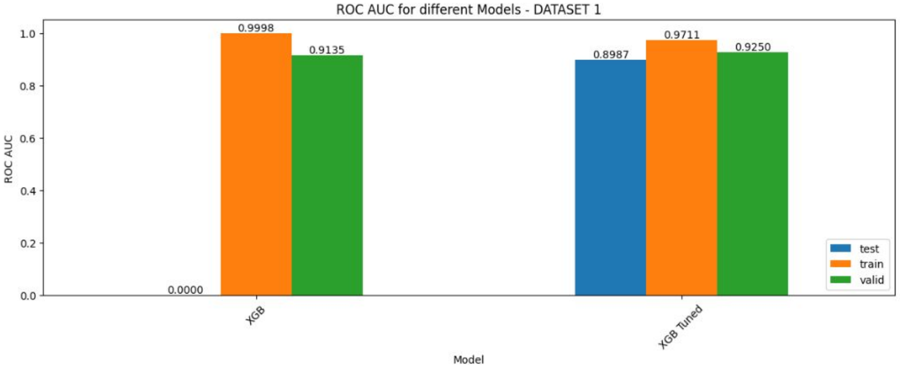


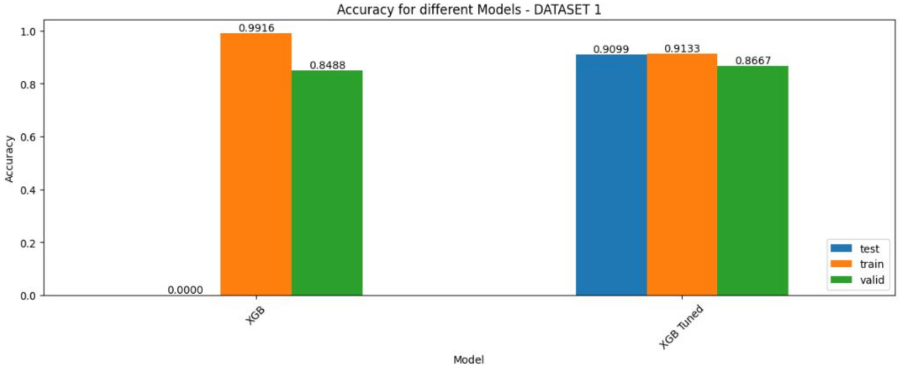


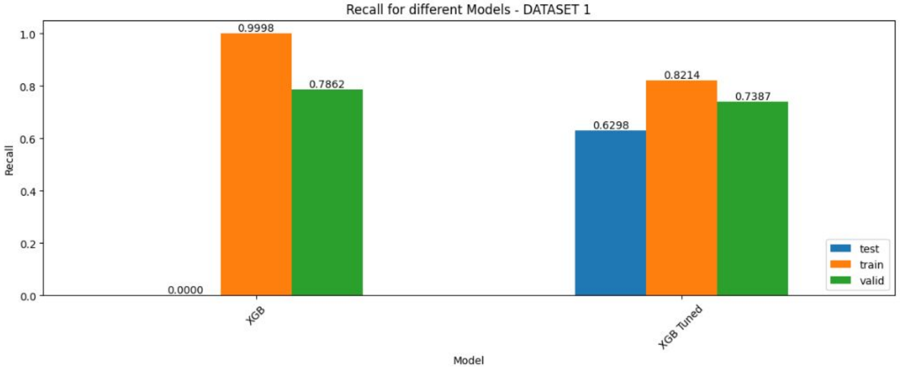


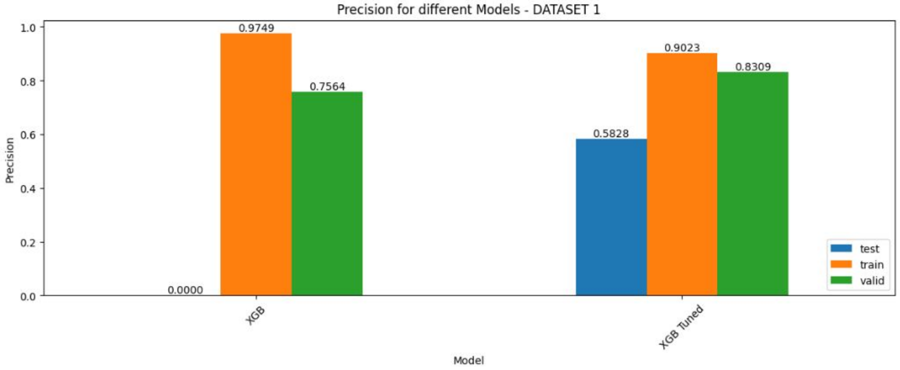
Riley

Classification Models – Hyperparameter Tuning

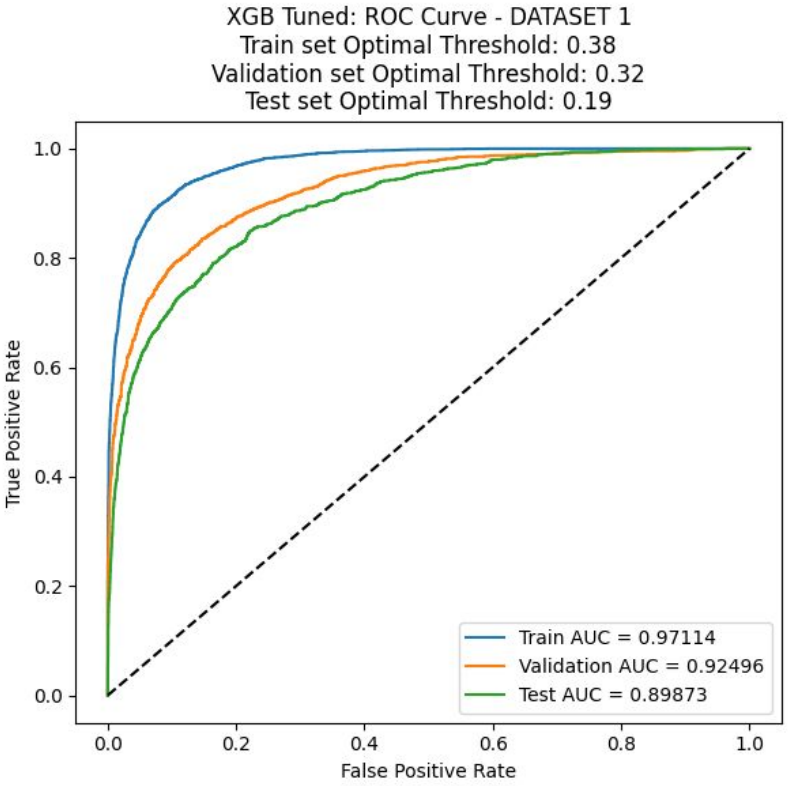








Classification Models – Hyperparameter Tuning



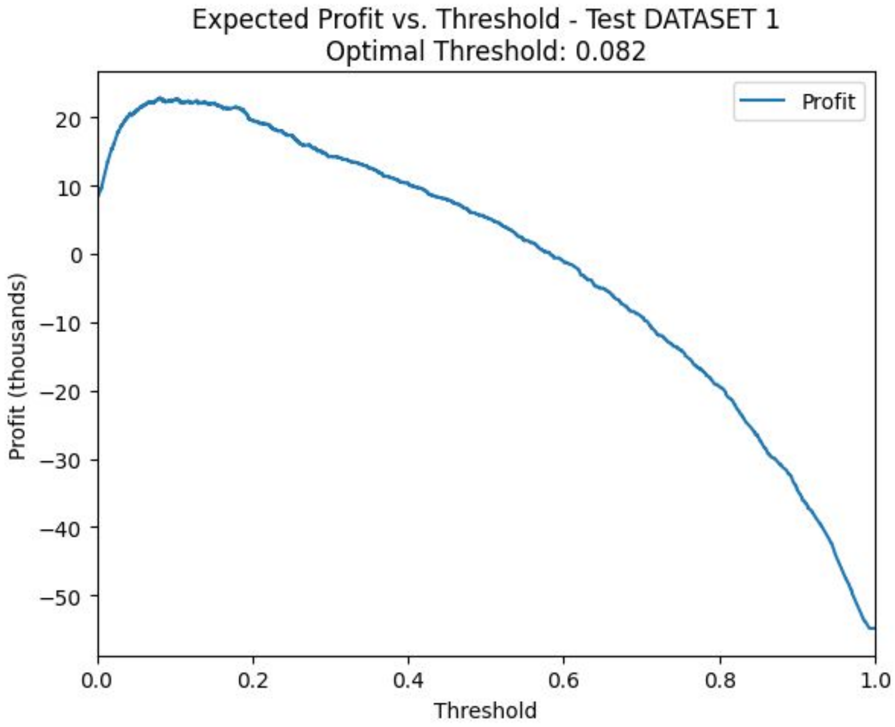
Expected Profit Calculation

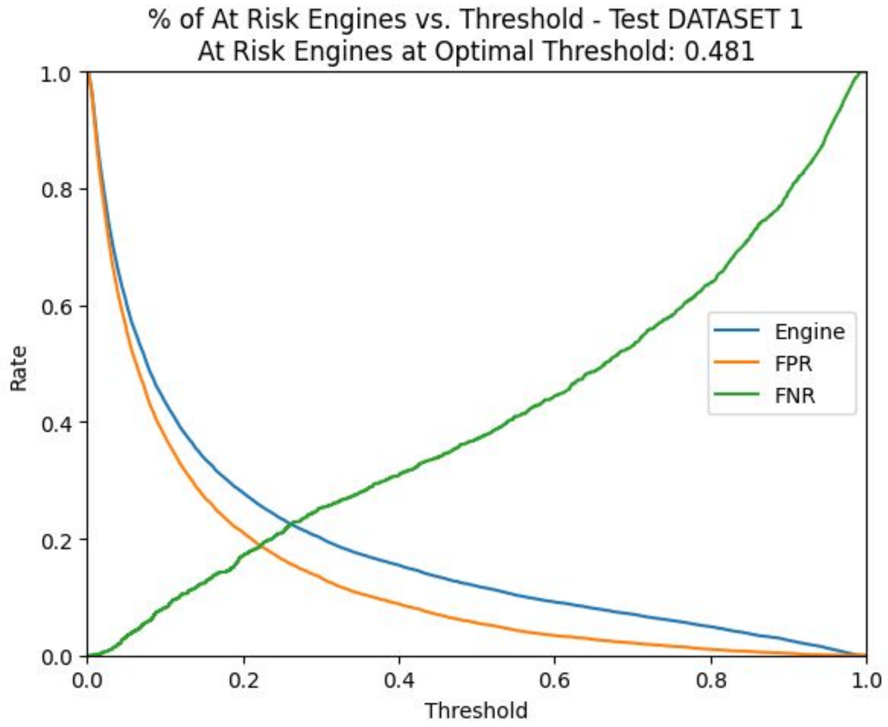
$$Expected\,Profit \ = \ P_P[TPR * B_{TP} \ + \ FNR * C_{FN}] \ + P_N[TNR * B_{TN} \ + \ FPR * C_{FP}]$$

Assume:

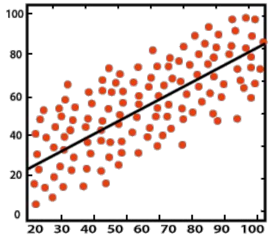
- TP Benefit = 400k
- TN Benefit = 0
- FN Cost = -500k
- FP Cost = -40k

Expected Profit Calculation

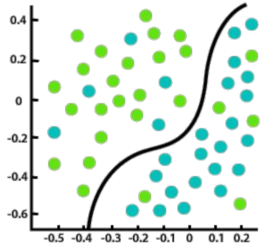




Recommendations



XGBoost Regressor



XGBoost Classifier

For equipment that requires **precise scheduling of maintenance** due to the greater care they require or their more expensive nature. Aramco can apply this to the following equipments:

1. Gas turbines and compressors
2. Drilling equipment
3. Pipeline monitoring systems

For the maintenance of less critical equipment, or for maintenance actions that have multiple components, by classifying them as ‘at risk’ or ‘not at risk’, **prioritising what actions they should take first** based on general thresholds rather than precise RUL predictions.

Conclusion